DEPARTMENT OF STATISTICS
Faculty of Applied Science & Technology
Islamic University
Kushtia-7003, Bangladesh

Syllabus for the Courses of M. Sc. in Statistics
Session: 2016– 2017; Final Examination: 2017

The courses for Master of Science (M. Sc.) program in Statistics promote at specializing and training in statistical methodology in its theoretical, practical/applied, and in scientific research aspects of modern age particularly with computer intensiveness. This program shall extend over one academic year (July to June). Students enrolled under this program comprise of two groups namely General Group (Group A) and Thesis Group (Group B). The courses offered by the department are of compulsory and optional nature. The total theoretical courses are 7, totaling 700 marks, 5 courses of which being compulsory courses and remaining 2 courses being optional courses. Students from each group have the option to choose any 2 of the available optional courses offered by the department.

DISTRIBUTION OF COURSES AND MARKS

The Total marks M. Sc program (General or Thesis group) shall be 1200 for all the subjects. The marks shall be distributed among the theoretical course. Practical, dissertation, Viva-Voce, project and field works.

The group-wise structure for the M.Sc. program is given below:

General Group (Group A)

Courses	No. of Courses	Total Marks
Theoretical	7	700
Practical	3	300
Project	1	100
Viva Voice	1	100
Total	**12**	**1200**

Thesis Group (Group B)

Courses	No. of Courses	Total Marks
Theoretical	7	700
Practical	1	100
Dissertation	2	200
Presentation on dissertation	1	100
Viva Voice	1	100
Total	**12**	**1200**

The examination of each theoretical course shall be of four hours duration and of 75 marks. A student shall answer 5 questions out of 8 while each question carries 15 marks. The remaining 25 marks are divided as 15 marks for Tutorial/Assignment and 10 marks for Class Attendance. The practical examination shall be of 12 hours duration for each course and of 80 marks. The remaining 20 marks are divided as 5 marks for Laboratory Performance/Assignment, 10 marks for Laboratory Attendance, and 5 marks for Laboratory Note Book.

Students shall have to appear before the *viva-voce* examinations at the end of the academic year.

ASSIGNMENT OF CREDITS

(a) One Class Hour per week for a period of 14 weeks shall be considered as one credit for each theoretical course, while Two Class Hours per week for a period of 14 weeks shall be considered as one credit for each practical course. A student shall have to complete at least 39 credits of the program. The credits of M. Sc degree program shall be as follows:

Credit per Course:	
Credit for each Theoretical course	= 3.5
Credit for each practical course/Field work /project	= 3.0
Credit for oral Examination	= 2.5
Total Credit for general group:	
Credit for Theoretical course	= 24.5
Credit for practical /Field work/project	= 12.0
Credit for Oral Examinations	= 2.5
	= 39.0
Total Credits for thesis group:	
Credit for Theoretical course	= 24.5
Credit for Comprehensive practical course	= 4.0
Credit for dissertation	= 5.0
Credit for presentation/viva-Voce for dissertation	= 3.0
Credit for Oral Examinations	= 2.5
Total Credits	= 39.0

The course distribution portraying code, title and their marks is given below:

Compulsory Courses

Course No	Course Title	*Marks*
STAT M1	Advanced Statistical Inference	100
STAT M2	Advanced Multivariate and Categorical Data Analysis	100
STAT M3	Sample Survey	100
STAT M4	Experimental Design	100
STAT M5	Survival Analysis	100

Optional Courses

Course No	Course Title	*Marks*
STAT. M6	Genomics and Bioinformatics	100
STAT. M7	Time Series Analysis and Forecasting	100
STAT M8	Econometric Analysis & Environmental Statistics	100
STAT M9	Probability Measure	100
STAT M10	Computer Intensive Statistics	100
STAT M11	Biometry and Genetic Statistics	100
STAT M12	Reliability Statistics and Modeling	100
STAT M13	Actuarial Statistics & Data Mining	100
STAT M14	Epidemiology and Population Studies	100

Practical Courses

General Group		
Course No	**Course Title**	**Marks**
STAT-LAB.M15	Statistical Data Analysis-I	100
STAT-LAB.M16	Statistical Data Analysis-II	100
STAT-LAB.M17	Statistical Data Analysis-III	100

Thesis Group

Course No	Course Title	Marks
STAT-LAB.M18	Statistical Data Analysis-IV	100

DETAILED SYLLABUS

STAT M1: Advanced Statistical Inference
(Full Marks: 100, 2 Lectures per week, Credit: 3.5)
(Year Final Exam: 75 Marks, 5 Questions out of 8 are to be answered)

Theory of Estimation:

Empirical Point Estimation:
Equivariance principle, EM Algorithm. Empirical Bayes (EB) Method, Bayes Conventional and Empirical Bayes Techniques, Approximation of Bayes and EB Method of Estimation for Prior Distribution, Application of EB Methods, Minimaxity and Admissibility in Exponential Families and their Properties, Bayesian Estimation in Linear Model, Predictive Inference with Reference to Bayesian Analysis, James Stein Estimator, Shrinkage Estimators.

Robust Statistic: Influence Functions, M-Estimators, L-Estimator, R-Estimator, Multidimensional Estimators, Application of Robust Estimators.

Method of Evaluating Estimators: U-Statistics and Its Properties, Best Unbiased Estimators and Their Properties.
Confidence Sets: Length of Confidence Intervals, Shortest Length Confidence Intervals, UMA and UMAU Confidence Sets, Randomized Confidence Sets, Invariant Confidence Sets.
Resampling Techniques: Review of Resampling Methods, Hypothesis Testing with Bootstrap, Cross Validation and other Estimates of Prediction, Adaptive Estimation and Calibration.
Simultaneous Confidence Interval: Bonferronis Method, Scheffe's Method in Linar Models, Tukey's Method in One-Way ANOVA Models, Confidence Bands for Cumulative Distribution Functions-
Non-parametric Estimation: Density estimation-introduction, histogram and series, naïve estimator, Idea of Kernel, Kernel, Fourier series, K-NN , idea wavelet, wavelet estimator and their MISE.

Regression function estimation- Kernel, orthogonal series and wavelet estimator- bandwidth selection-cross validation method

Theory of Hypothesis Testing:

Optimal Test, Locally Best Test, Tests Under Restricted Alternatives, Similar Region and Neyman Structure, Most Powerful Similar Region (MPSR) Test, Uniformly most Powerful Similar Region (UMPSR) Test, Asymptotic Efficiency of Test, Sequential Probability Ratio Test (SPRT) for Three Hypotheses, Sobel and Wald Test, Lagrange Multiplier (LM) Test, Test in Presence of Nuisance Parameters, Union-Intersection and Intersection-Union Test, Armitage Method for Composite Hypotheses, Sequential T, χ^2 and T^2 test.

Empirical Bayes Testing of Multiple Hypothesis, Lindley's Procedure for Test of Significance, Lindley's Paradox, p-value and Bayesian Significance Probability, Bayes in Test in Linear Model.

Texts:

1. George, C. and Berger, R. L. (2003): *Statistical Inference,* 2nd edition, Thompson-Duxbury, USA.
2. Hogg, R. H., Mckean, J. W. and Craig, A. T. (2007): *Introduction to Mathematical Statistics*, 6th edition, Pearson Education (Singapore) Pte Ltd.
3. Lehman, E. and Cassela, G. (1998): *Theory of Point Estimation*, Springer Verlag, New York.

References:

1. Lehman, E. L. (1997): *Testing Statistical Hypothesis,* 2nd edition, Springer-Verlag, New York.
2. Rohatgi, U. K. and Saleh, A. K. Md. E. (2005): *An Introduction to the Probability and Statistics,* 2nd editions, John Wiley and Sons Inc., New York.
3. Bansal, A. K. (2007): *Bayesian Parametric Inference*, Norosa Publishing House, India.
4. O' Hagan, A and Forster, J. (2004): *Advanced Theory of Statistics, Bayesian Inferences*, Vol. 2B, Arnold.
5. Carlin, B. P. and Louis, T. A. (2002): *Bayes and Empirical Bayes Method for Data Analysis* 2nd edition, CRC Press.
6. Stuart, A. and Keith, J. O. (1986): *Advanced Theory of Statistics*, Vol. II, Charles Griffin and Company Ltd. London.
7. Kalbfleisch, J. G. (1985): *Probability and Statistical Inference,* Vol. 1 and Vol. II. 2nd edition, Springer-Verlag, New York.
8. Wasan, M. T. (1970) *Parametric Estimation*, McGraw-Hill Inc, New York.
9. Knight, K. (2000): *Mathematical Statistics*, Chapman & Hall, London.
10. Shao, J. (1999): *Mathematical Statistics*, Springer-Verlag, New York
11. Mukhopadhyaya, P. (1996): *Mathematical Statistics*, New Central Book Agency (P) Ltd, India.

STAT M2: Advanced Multivariate Analysis and Categorical Data Analysis
(Full Marks: 100, 2 Lectures per week, Credit: 3.5)
(Year Final Exam: 75 Marks, 5 Questions out of 8 are to be answered)

Theory of Distance and Its Application to Classification Problems: Objectives of Multivariate Analysis, Concept of Distance in Multivariate Analysis, Different Measures of Distance, Some Properties of Distance Function, Multivariate Normal Distribution, Assessing Assumptions of Multivarite Normality by Box-Cox Transformation, Detecting and Cleaning Outliers, Transformation of Multivariate Normal Observations to Near Multivariate Normal, Inference About Mean Vector, Hotteling T^2, Confidence Region, Conferring Methods of Multiple Comparison.

Analyzing Association among Variables: Measuring and Interpreting Association Especially Between Two Variables, Graphical Investigation of Many Associations, Correction of Correlation for Effects of Extraneous Variables, Measuring Association Between Two Sets of Variables, Testing Hypotheses About Sets of Associations, Test that All Population Correlations are Zero (Mutual Independence of All Variables), Test that All Population Canonical Correlations are Zero, Test that Some Population Canonical Correlations are Zero.

Two and Three-Way Contingency Tables: Log-Linear Models, Interpretation of Log-Linear Parameters, Choice of Model, Detection of Model Deviations, Log-Linear Hypotheses, Estimation, Testing Hypotheses.

Multi-Dimensional Contingency Tables: Log-Linear Models, Classification and Interpretation of Log-Linear Models, Choice of Model, Diagnostics, Model Search Strategies.

Analysis of Covariance Structure: Covariance Structure, Hypotheses about Covariance Structure, Model of Covariance Structure Analysis, Scope of Covariance Structural Analysis, Illustration of Likelihood Ratio Test, Illustration of Likelihood Ratio Test Procedures, Tests of Covariance Structure Based on Union-Intersection Principle, Illustration of Structural Analysis Based on Union-Intersection Principle, Sphericity Test, Advantages of Covariance Structure Analysis, Assumptions of Covariance Structure in Statistical Analysis, Estimation of Variance Components, Confirmatory Factor Aspect of Multivariate Analysis, Power of Statistical Tests When Covariance Structure is Known, Tests for Broad Class of Covariance Structures, Bayesian Inference in Multivariate Regression, Multivariate Analysis of Covariance, Checking and Violation of

Assumptions, Two-way Multivariate Analysis of Variance (MANOVA), Profile Analysis, Repeated Measures With Growth Curves.

Linear Discriminant Functions: Introduction, General Ideas, Linear Programming and Perception Criterion, Error Correction and Perception Criterion, Fisher's Criterion, Least Squares Approach, Choosing an Estimation Method, Specializations and Generalizations.

Cluster Analysis: Distance Measures, Hierarchical Methods, Agglomerative Methods, Divisive Methods, Optimization Criteria, Optimization Algorithms, Method of Mixtures, Wishart's Mode Method, Clumping Techniques.

Factor Analysis and Latent Structure Analysis: Factor Analysis Models, Estimation, Tests for Appropriate Number of Factors, Interpretation of Factor Analysis, Rotation Problem of Factor Analysis, Factor Analysis and Principal Components, Latent Structure Analysis.

Independent Component Analysis (ICA): Basic Concept on ICA: Definition, Identifiability of the ICA Model, Ambiguities of ICA, Preprocessing for ICA: Centering, Whitening, Principles of ICA Estimation: Maximization of non-Gaussianity using kurtosis and Negentropy, Minimization of Mutual Information, Maximum Likelihood Estimation.

Multivariate Multiple Linear Regression Model: Test for Regression Parameters, Prediction from Multivariate Multiple Regression, Prediction of Survival Variables, Comparing Two Forms, Mean Correlated Form of Regression Model, Multiple Regression Model with Time Dependent Errors.

Texts:

1. Anderson, T. W. (1984): *An Introduction to Multivariate Statistical Analysis*, 2nd edition, John Wiley, New York.
2. Agresti, A. (2002): *Categorical Data Analysis*, 2nd edition, John Wiley, New York.

References:

1. Jonhnson, R. A. and Wichern, D. W. (2002): *Applied Multivariate Statistical Analysis*, 7th edition, John Wiley, New York.
2. Mardia, K. V., Kent, J. T., and Bibby, J. M. (1979): *Multivariate Analysis,* 3rd edition, Academic Press, London.
3. Hyv̈arinen , A. Karunen, J. and Oja, E. (2001): *Independent Component Analysis*, New York: Wiley.

STAT M3: Sample Survey

(Full Marks: 100, 2 Lectures per week, Credit: 3.5)

(Year Final Exam: 75 Marks, 5 Questions out of 8 are to be answered)

Review and Application of Multi-Stage Sampling, General Approach to Obtain Estimators. Their Variances, and Estimated Variances, Optimum Sampling and Sub-Sampling Fractions Considering Cost.

Review and Application of Multiphase Sampling Technique, Repeated Sampling of Some Populations, Longitudinal Surveys.

Some Problems of Inference under Fixed Population Set Up, Inference from Finite Population Using Prediction-Theoretic Approach, Prediction under Multiple Regression Model, Inference from Complex Samples.

Various Techniques Used in Determining Sample Size Under Different Sampling Techniques, Estimating Population Size Using Direct Inverse Sampling, Estimating Population Density, Design Effects and Methods for their Treatment.

Sources and Types of Survey Errors, Non-Response, Types and Its Effects on Accuracy of Estimators, Techniques for Adjustment of Non-Response, Spatial Sampling, Spatial Prediction or Kriging, Spatial Designs, Plot Shapes.

Adapting Sampling, Adaptive Sampling Design, Adaptive Cluster Sampling, Stratified Adaptive Cluster Sampling.

Non-sampling Errors, Modified Hansen-Hurwitz Estimator, Randomized Response Method for Quantitative Data, Deming's Model for Effects of Call-backs, Politz-Simmons Technique, Randomized Response Method.

Methods of Increasing Response Rates, Errors of Measurement, Mathematical Models, Treatment of Missing Data, Imputations, Multiple Imputation for Non-Response, Maximum Likelihood Method for Non-Response.

Selected Topics in Sample Design and Estimation Methodology, WHO-EPI Survey, Integrated Multipurpose Probability Sample (IMPS), Bangladesh Demographic and Health Survey (BDHS), Labour Force Survey(LFS), Bangladesh Maternal Mortality Survey (BMMS), Household Income Expenditure Survey, Poverty Monitoring Survey, SVRS.

Texts:

1.Thompson, S. K. (1992): *Sampling*, John Wiley, New York.

2.Raj, D. and Chandok, P. (1998): *Sample Survey Theory*, Narosa Publishing House, India.

References:

1. Cochran, W. G. (1977): *Sampling Techniques*, 3rd edition, Wiley, New York.
2. Sukhatme, P. V. and Sukhatme, B. V. (1977): *Sampling Theories of Surveys with Applications*, ISAS, New Delhi.
3. Kish, L. (1995): *Survey Sampling,* Wiley, New York.
4. Chudhury, A. and Stenger, H. (1992): *Survey Sampling Theory and Methods*, Chapman and Hall Tayor and Francis Group, Yew York.
5. Levey, P. and Lemeshow, S. (1999): *Sampling of Population: Methods and Application*, Wiley, New York.
6. Mukhapadhyay, P. (1998): *Theory and Methods of Survey Sampling,* Prentice Hall, New Delhi.
7. Sharon L. Lohr (2010), *Sampling: Design and Analysis*, Second Edition, Arizona State University, Brooks/Cole, Cengage Learning

STAT M4: Experimental Design

(Full Marks: 100, 2 Lectures per week, Credit: 3.5)

(Year Final Exam: 75 Marks, 5 Questions out of 8 are to be answered)

Linear Estimation, Linear Hypothesis, General Principles and Practical Difficulties in Controlled Experiments, Strategies in Planning Experimental Programs.

Higher level Factorial Experiments: Concept of linear, quadratic, cubic, etc. with main and interaction effects of S^n factorial experiment, estimation of different effects and analysis procedures, Problems of f.e., Different types of confounding, confounding with factorial experiments of level greater than 2, their estimation of effects and analysis. Methods of simultaneous confounding. Detection of factorial effects in confounding factorial experiments.

Fractional Factorial Design: Basic concepts, alias, defining contrast, etc., fractional replication of the S^n and S^{n-p} factorial design and their analysis.

Multifactor Experiments with Randomization Restrictions: Randomized Blocks and Latin Squares as multifactor designs, Split plot design, Split block design. related problems.

Incomplete Block Design: Main concept, Comparison with complete designs. Relationship between parameters and their proofs, analysis of BIBD (intra-block and inter-block analysis), combined estimate, missing value estimate, use of orthogonal Latin Squares and construction of BIBD, PBIBD, Lattice design, Youden Squares, Galois field and finite projective geometry with application.

Response Surface Methodology: Main concepts, the method of Steepest Ascent, analysis of quadratic models, Response surface designs, designs of fitting first and second order model, Evolutionary operation, related problems.

Some advanced designs: Weighting designs, Repeated measurement designs, optimal designs, groups of experiments, etc.

Text:

1. Montegomery, D. C. (2003): *Design and Analysis of Experiments*, 5th edition, John Wiley, New York.

References:

1. Das, M. N. and Giri, N. C. (1997): *Design and Analysis of Experiments*, 2nd edition, New International (P) Ltd., India
2. Kempthorne, O. (1952): *The Design and Analysis of Experiments*, Wiley, New York.
3. Graybill, F. A. (1961): *An Introduction to Linear Statistical Models*, Vol. I, McGrow-Hill, New York.
4. Federer, W. T. (1955): *Experimental Design: Theory and Application*, Macmillan, New York.
5. Cochran, W. G. and Cox, D. R. (2000): *Experimental Design,* 2nd edition, John Wiley, New Delhi.
6. John, P. W. M. (1971): *Statistical Design and Analysis of Experiments,* Wiley, New York.

STAT M5: Survival Analysis

(Full Marks: 100, 2 Lectures per week, Credit: 3.5)

(Year Final Exam: 75 Marks, 5 Questions out of 8 are to be answered)

Review of parametric, nonparametric survival models, Proportional hazard model and censored data.

Accelerate failure time models: General and Parametric accelerate failure time models, fitting and comparing accelerated failure time models.

Stratified Cox Procedure: General Stratified Cox model, No-Interaction Assumptions and its test, Several Stratification Variables, Graphical View of the Stratified Cox Approach, Stratified Cox Likelihood.

Extension of the Cox Proportional Hazards Model for Time-Dependent Variables: Review of the Cox PH Model, Definition and Examples of Time-Dependent Variables, Extension of the Cox Proportional Hazards Model for Time-Dependent Variables: The Extended Cox Model for Time-Dependent Variables, Hazard Ratio for the Extended Cox Model, Assessing Time-Independent Variables, Application of the Extended Cox Model to An Epidemiologic Study, Extended Cox Likelihood, Truncated, Left censored and interval censored data.

Identification of Risk Factors Related to Dichotomous and Polychotomous Outcomes: Univariate Analysis, Logistic and Conditional Logistic Regression Models for Dichotomous Responses, Multinomial Logistic Regression for Polychotomous Outcomes.

Recurrent Event Survival Analysis: Robust Estimation, Parametric Approach Using Shared Frailty, Survival Curves with Recurrent Events

Competing Risks Survival Analysis: Overview, Examples of Competing Risks Data, Sweparate Models and Lunn–McNeil (LMalt) Approach of analyzing competing risks,

Cumulative Incidence Curves (CIC), Conditional Probability Curves (CPC).

Design Issues for Randomized Trials: Background to Statistical Decisions, Design Considerations for Time-to-Event Outcomes, Determining the Required Number of Events (NEV), Determining the Required Total Number of Participants, Other Design Issues.

Loglinear Models for Contingency Tables: Loglinear Models for Two-Way Tables, Loglinear Models for Independence and Interaction of Three-Way Tables, Loglinear Models for Higher Dimensions, The Loglinear Logit Model Connection, Loglinear Model Fitting: Likelihood

Equations and Asymptotic Distributions, Loglinear Model Fitting: Iterative Methods and their Application.

References

1. Agresti, A. (2002). Categorical Data Analysis, 2nd edition Wiley.

2. Kleinbaum, D. G. and Klein, M. (2012), Survival Analysis: A Self Learning Text, 3rd edn, Springer.

3. Lawless, J. F. (2003): *Statistical Models and Methods for Life Time Data*, 2nd edition, Wiley Series, New York.
4. Lee, E. T. (1992): *Statistical Methods for Survival Data Analysis*, 2nd edition, Wiley Series, New York.
5. Kalbfleisch, J. D. and Prentice, R. L. (1980): *The Statistical Analysis of Failure Time Data*, John Wiley, New York

6. Fienberg, S. F. (1980): *The Analysis of Cross-Classified Categorical Data*, 2nd edition, The MTT Press.
7. McCullah, P. and Nelder, J. A. (1982): *Generalized Linear Models*, Chapman and Hall, London.

8. Daniel, W. W. (2000): *Bio-Statistics: A Foundation for Analysis in the Health Science*, 7th edition, John Wiley and Sons, New York.
9. Kenneth, J. R. and Greenland, S. (1998): *Modern Epidemiology,* 2nd edition, Lippincolt Williams and Wilkins.
10. Indrayan, A. and Sarmukaddam, S. B. (2001): *Medical Biostatistics*, CRS Press.
Newman, S. (2001): *Biostatistical Methods in Epidemiology*, Wiley, New York.
11. Sahai, H. and Khurshid, A. (1995): *Statistics in Epidemiology, Methods, Techniques and Applications*, CRC Press, London.

STAT M6: Genomics and Bioinformatics

(Full Marks: 100, 2 Lectures per week, Credit: 3.5)

(Year Final Exam: 75 Marks, 5 Questions out of 8 are to be answered)

Basic Genomics: Introduction. Classification of Genomics. Genes and Chromosomes. Cell division, Nucleic acid, Molecular Genetics (Marker, DNA sequence, SNPs). Genotype and Genotyping technology. Central dogma (DNA/Gene transcription, RNA translation and Protein synthesis).

DNA Sequence Analysis: DNA sequencing. Classification of sequencers including NGS. Analysis of DNA patterns, Overlaps counted, Overlaps not counted and motifs, Sequence accuracy, Sequence formats, Conversions of one sequence format to another. Single and multiple sequence alignments approaches. Phylogenetic analysis DNA Sequence. Some bioinformatics databases including GeneBank, NCBI, PDB, BLAST and FASTA.

QTL Analysis: .Introduction. Marker Analysis of Phenotypes. Whole-Genome Marker Analysis. The Structure of QTL Mapping (Population and Quantitative Genetic Structure of the Mixture Model). Interval Mapping Approaches for QTL Analysis (Linear regression and maximum likelihood approaches for QTL analysis with backcross and F2 populations). Composite and multiple interval mapping approaches for QTL analysis.

Gene-Expression Data Analysis: Introduction to different types of microarray gene expression data. Preprocessing (Transformation, normalization, Image analysis and filtering). Identification of differential expressed (DE) genes in two or more groups using statistical test. Clustering and Classification for Gene-Expression Data Analysis. Inferring genetic regulatory networks from microarray experiment with Bayesian networks. Modeling genetic regulatory networks using gene expression profile. Gene-set enrichment analysis.

Genome-wide Association Studies (GWAS): Introduction . QTL and SNP analysis with Gene-Expression data. SNP analysis using contingency table. GWAS using linear mixed models and GeneABLE. Haplotype Estimation. Regional multilocus association models. Linkage disequilibrium and tagging. Practical guide to linkage disequilibrium analysis and tagging using Haploview.

Main Books:

Ben Hui Liu and Leming M Shi. (2013). *Statistical Genomics and Bioinformatics***,** Chapman and Hall/CRC press, 2nd edition, New York.

David W. Mount. (2004). *Bioinformatics: Sequence and Genome Analysis*, Second Edition, Cold Spring Harbor Laboratory Press.

Recommended Books:

Xu, S. (2013). *Principles of statistical genomics*. Springer.

Liu, B. H. (1997). *Statistical genomics: linkage, mapping, and QTL analysis*. CRC press.

Ferreira, M. A. R., Medland, S. E., & Posthuma, D. (2008). *Statistical genetics: gene mapping through linkage and association*. New York: Taylor & Francis.

Gondro, C., Van der Werf, J., & Hayes, B. J. (Eds.). (2013). *Genome-wide Association Studies and Genomic Prediction*. Humana Press.

STAT M7: Time Series Analysis and Forecasting

(Full Marks: 100, 2 Lectures per week, Credit: 3.5)

(Year Final Exam: 75 Marks, 5 Questions out of 8 are to be answered)

Time Series Analysis

Stationary ARIMA processes: Stationary Ergodicity, White noise, Moving average processes, Autoregressive processes, ARIMA processes, The autocovariance – generating functions.

Non stationary time series: Linear time trend, unit root, comparison of trend-stationary and unit root processes, Hypotheses testing for the simple time trend model. Bayesian analysis and unit roots, VAR containing unit roots, Spurious regression.

Vector processes: Vector autoregression, auto covariances results for vector processes, The spectrum for vector processes, ML estimation and hypothesis testing for an unrestricted VAR. bivariate granger casualty tests. ML estimator of restricted VAR, the impulse- response function, structural econometric models,

Cointegration: Introduction, testing the hypotheses of no cointegration, testing hypotheses about the cointegrating vector, FIML analysis of cointegrated systems, canonical correlation.

Non-linear time series: Bilinear models, Autoregressive conditional heteroscedasticity (ARCH) model, Detection of ARCH, Estimation under ARCH, Generalized autoregressive conditional heteroscedasticity (GARCH) model, Smooth transition (threshold) autoregressive (STAR) model.

Forecasting

Principles of forecasting: Preliminaries, Prediction with a population regression function (PRF), Forecasting with a sample regression function (SRF), Simple extrapolation model, Smoothing and extrapolation time series, Forecast based on finite and infinite number of observations, Linear projection, Optimal forecasts for Gaussian processes, Box-Jenkins modeling philosophy.

Forecasting with ARIMA and other models: Minimum mean square error forecasts, Computing a forecast, The forecast error, Forecast confidence intervals, Properties of ARIMA forecasts. Forecasting with seasonal ARIMA (SARIMA) model, Forecasting with multivariate ARIMA (MARIMA) model.

Forecasting techniques: ARAR algorithm, Holt-Winters algorithm, Holt-Winters seasonal algorithm, Choosing a forecasting algorithm.

Building dynamic regression models and intervention analysis: The linear transfer function (LTF), Identification of LTF. Forecasting with LTF and other transfer functions, Pulse intervention, step intervention, Building intervention models, Multiple and compound interventions.

Intervention and outlier detection: The rationale for intervention and outlier detection, Models for intervention and outlier detection, Likelihood ratio criteria, An iterative detection procedure, Detected events near the end of the series.

References:

1. Box, G.E.P. and Jenkins, G.M. (1976): Time Series Analysis: Forecasting and Control, Holden-Day, Sun Francisco.
2. Hamilton, J.D. (1994): Time Series Analysis, Princeton University Press, N.J.
3. Pankratz, A. (1991): Forecasting with Dynamic Regression Models, Wiley, N.Y.
4. Pindyck, R.S. and Rubinfeld, D.L. (1998): Econometric Models and Econometric Forecasts, 4th ed., Irwin-McGraw Hill, N.Y.
5. Books Recommended:
6. Anderson, T.W. (1971): The Statistical Analysis of Time Series, Wiley, N.Y
7. Barnett, W. et al. (1990): Non parametric and Semi-parametric Methods in Economics and Statistics, CUP, N.Y.
8. Engle, R and McFadden, D (eds) (1993): Handbook of Econometrics, Vol.1-4 North Holland, N.Y.
9. Findley, D.F. (1981): Applied Time Series Vol. I & II, Academic press, N.Y.
10. Fuller, W.A (1976): Introduction to Statistical Time Series, Willey N.Y
11. Giannini, C (1992): Topic in Structural VAR Econometrics, Springer-Verlag.
12. Zellner, A. 1971): An Introduction to Bayesian Inference in Econometrics, Willey, N.Y.

STAT M8: Econometric Analysis & Environmental Statistics

(Full Marks: 100 (Part A 70 Marks & Part B 30 Marks),

2 Lectures per week, Credit: 3.5)

(Year Final Exam: 75 Marks, 5 Questions out of 8 are to be answered)

Part A

Econometric Analysis

SEM, Fundamentals, Identification Problem, Estimation Method, Method of Estimation, Specification Tests.

Indirect Least Squares (ILS), Two Stage Least Squares (2SLS), Three Stage Least Squares (3SLS), Full Information Maximum Likelihood (FIML), Seemingly Unrelated Regression Equation (SURE), Monte-Carlo Studies, Concept of Model Selection in Econometrics.

Estimation, Testing and Predicting Panel Data Model With Individual and Time Effect Model.

Non-linear Models in Econometrics, Bayesian Inference in Econometrics.

Optimization in Econometrics, Functional Form, Non-linearity and Specification, Dummy Variables, Non-linearity in Variables, Specification Analysis, Biased Estimators, Functional Form, Non-linearity and Specification, Biased and Pre-test Estimators.

Non-linear Regression Equation, Non-linear Regression Models, Estimators of Non-linear Regression Models, Hypothesis Testing and Parametric Restrictions, Box-Cox Transformation.

System of Regression Equations, Covariance Structure, Random Coefficient Model, Seemingly Unrelated Regression Model, Non-linear System and GMM Estimation.

Non-spherical Disturbances, Generalized Regression and GMM Estimation and Generalized Estimating Equation (GEE), Consequences for LSE, GLS, MLE and GMM Estimation, Testing Hypothesis in GMM Framework Models With Discrete Dependent Variable, Models for Panel Data for Fixed Random Effects, Autocorrelation and Dynamic Model.

Klein and Goldberger Model, Wharton Group of Models, Brookings Quarterly Econometric Model, FRB ST. Louis Model, Liu and Hwa Monthly Model, Trace Model, RDX2 Model, Policy Appraisal by Econometric Models, Models of Disequilibrium, Tests of Equilibrium.

Texts:

1. Greene, W. H. (2012): *Econometric Analysis*, 7th edition, Prentice-Hall, India.
2. Sukesh K. Ghosh (1991): *Econometrics Theory an Applications*, Prentice-Hall, Inc., USA.

References:

1. Judge, G. G. (1988): *Introduction to the Theory and Practice of Econometrics*, 2nd edition, John Wiley and Sons, New York.
2. Gujarati, D. (2012): *Basic Econometrics*, 5th ed. McGraw-Hill
3. Johnston, J. (1977): *Econometric Methods* (4th ed.), McGraw –Hill, New York.
4. Koutsoyiannis, A. (1977): *Theory of Econometrics* (2nd ed.)
5. Johnston, J. and Dinardo, J. (1997): *Econometric Methods*, 4th ed. The McGraw-Hill Com.

Part B

Environmental Statistics

Understanding Environmental Pollution: Pollution and its importance, Why does pollution happen? Pollutant sources, Detailed study on Air and Water pollution; Global climate change.

Stochastic processes in the environment: Applications of Bernoulli, Poisson and normal Processes to Environmental Problems.
Environmental Sampling: Network Sampling, Composite Sampling, Ranked-set Sampling.

Detectability of Sampling: Basic Concept of Detectability, Constant Detectability Over a Region, Estimating Detectability, The Effect of Estimated Detectability, Detectability with Simple Random Sampling.

Diffusion and Dispersion of Pollutants: Wedge Machine, Particle Frame Machine and Plume Model.

Dilution of Pollutants: Deterministic Dilution, Stochastic Dilution, and the Theory of Successive Random Dilution (SRD), Applications of SRD to Environmental Phenomena: Air Quality, Indoor Air Quality, Water Quality, Concentrations of Pollutants in Soils, Plants and Animals and Concentrations in Foods and Human Tissue.

Statistical Theory of Rollback (STR): Predicting Concentrations After Source Control, Correlation, Previous Rollback Concepts, Environmental Transport Models in Air and Water.

Text:

1. Barnett, V. and Turkman, K. F. (1993): Statistics for the Environment. John Wiley & Sons, Ltd., Chichester.

References:

1. Barnett, V. (2004): Environmental Statistics: Methods and Apllications. John Wiley & Sons, Ltd.
2. Thomson, S. (1992): Sampling. John Wiley & Sons Inc..
3. Wayne, R. Ott. (2000): Environmental Statistics and Data Analysis. Lewis Publishers.
4. Hill, M. K. (2004): Understanding Environmental Pollutions, Combridge University Press

STAT M9: Probability Measure

(Full Marks: 100, 2 Lectures per week, Credit: 3.5)

(Year Final Exam: 75 Marks, 5 Questions out of 8 are to be answered)

Sets and Classes of Events: Algebra of sets, Relations, Open and closed set on R^n. The events and classes of events.

Measure: σ-Algebra, Measurable set, The concept of measurability, Lebesgue measure on the real line, Elementary properties of measures, Borel set.

Random Variables and Function: Random variable, Limit of Random variables, Function, Inverse function, Measurable function, Simple function, Borell function, Characteristic function, Random variable as measurable functions.

Integral of Measureable Function: Lebesgue integral of simple functions, Integrable functions, sequences of integrable functions, The general and the Reimann-Stieltje's integral.

Probability Measure: Definition of probability, Some simple properties, Discrete probability space, General probability space, Induced probability space, Extension of probability, Probability measure, Lebesgue-Stieltje's measure, Signed measure, Borel-Cantellilemmas, Zero-one Law, Kolmogorov's Zero-one law.

Distribution Functions and Expectations: Distribution function of a random variable and a random vector, Decomposition of D.F's, Correspondence theorem, Definition of expectation, Properties of expectation, Moments and inequalities.

Convergence of Random Variable and Distribution: Convergence in probability, Convergence almost surely, Convergence in rth mean, Convergences in distribution and their relations, Convergence of distribution functions and characteristic functions, Convergence of moments.

References:

1. Bhat, B.R. (1988): Modern probability theory, Wiley Eastern.
2. Billingsley, P. (1995): Probability and Measure, Wiley, N.Y.
3. Kingman, J.F.C and S.J. Taylor (1997): Introduction to Measure and Probability, C.U.P., London.
4. Dudley (1994): Real Analysis, McGraw-Hill, N.Y.
5. Gnedenko, B.V. and A. N. Kolmogorov (1984): Limit Distribution for Sums of Independent Random Variables. Addison-Wesley. N.Y.
6. Halmos, P.R. (1974): Measure Theory; Spring-Verlag, N.Y.
7. Loeve, M. (1960): Probability Theory, Wiley, N.Y.
8. Pitt, H.R (1985): Integration Measure and Probability, Oliver and Boyd.
9. Rudin, W. (1994): Real and Complex Analysis, McGraw-Hill, N.Y.

STAT M10: Computer Intensive Statistics

(Full Marks: 100, 2 Lectures per week, Credit: 3.5)

(Year Final Exam: 75 Marks, 5 Questions out of 8 are to be answered)

Jackknifing: Bias Correction, Pseudo-Value, Approximate Confidence Intervals, Extension to 2-or More-Sample Problems.

Bootstrapping: Bootstrap Strategy, Sampling Distributions, Empirical Distributions, Bootstrap Distributions, Percentile Bootstrap: Definition and Use of Confidence Limits, Relation to Jackknife, Application to Hypothesis Testing, Number of Simulation Samples Required, Variants: Smoothed Bootstrap, Bias-Corrected Bootstrap, Computational Aspects, Balanced Re-sampling.

Simulation Testing: Randomization Test, Approximate Randomization Test, Monte-Carlo Tests, Examples, Unbiasedness, Power, Number of Simulated Samples Needed.

Density Estimation: Definition, Examples, Bias, MSE and IMSE, Choice of Kernel and Smoothing Parameter, Computation Via Fast Fourier Transform, Kernel Estimates for Non-Negative and Circular Data, Variable and Adaptive Kernel Estimators, Kernel Estimation in Nonparametric Regression.

References:

1. Efron, B. and Tibshirani, R. J. (1993): *An Introduction to Bootstrap,* Chapman and Hall, New York.
2. Efron, B. (1984): The Jackknife, the Bootstrap and Other Re-sampling Plans, Society for Industrial Mathematics.

3. Noreen, E. W. (1982): *Computer-Intensive Methods for Testing Hypothesis*, Wiley, New York.
4. Shao, J. and Tu, D. (2000): *Jackknife and Bootstraps,* Springer-Verlag, New York.
5. Silverman, B. W. (1986): *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.

STAT M11: Biometry and Genetic Statistics

(Full Marks: 100 (Part A 50 Marks & Part B 50 Marks),

2 Lectures per week, Credit: 3.5)

(Year Final Exam: 75 Marks, 5 Questions out of 8 are to be answered)

Part A

Biometry

Biometry and Biostatistics. Biological variability Bioassay-meaning, design for bioassay; direct, indirect and parallel line assays. Slope ratio assays. The tolerance distribution. Estimation of effective dose, relative potency, natural response, association between disease and exposure, confounding and interaction case control studies, matched case-control studies statistical analysis-study design, odds ratio, attributable and relative risks. linear and non linear logistic regression model, log-normal model. Application of life table method to biological population. genetic models. incomplete block design for bio-assays.

Part B

Genetic Statistics

Description of genetic mechanism. Reproduction and the hardy Weinberg Law. Random mating. Equilibrium in randomly mating population. Gamete and Genotype distributions. Phenotype distributions. Inbreeding and non-random mating. Assortative mating. Genetic linkage. X- linked genes. Selection, mutation and migration. Distribution of mutation. Joint effects of matuation and selection, selfing. Self sterility alleles. Stationary distributions of gene frequencies. Estimation of genetic parameters. Test of genetic hypothesis.

References:

1. Elandt Johnson, R.C. (1957): Probability Models and Statistical in Methods in Genetics, John Wiley, N.Y.
2. Falconer, D.S. (1986): Introduction to Quantitative Genetics, Oliver and Boyd.
3. Fraser Roberts, J.A. & Pemlercy, N.E. (1978): An Introduction to Medical Genetics. Oxford University Press.
4. Kempthorne, O. (1957): Introduction to Genetic Statistics. John Wiley, N.Y
5. Li, C.C. (1976): Population Genetics. Chicago U.P.

STAT M12: Reliability Statistics and Modeling

(Full Marks: 100, 2 Lectures per week, Credit: 3.5)

(Year Final Exam: 75 Marks, 5 Questions out of 8 are to be answered)

Reliability and quality concepts and models for reliability data: Concepts of quality and reliability. Examples and features of reliability data. Strategy for data collection, modeling, and analysis of reliability data. Models for continuous failure-time processes. Models for discrete data from a continuous process. Quality assurance. Concept of total quality management (TQM), Cause-and-effect-diagrams. Failure modes and effect analysis (FMEA). Failure modes effect critically analysis (FMECA).

Reliability estimation: Non-parametric estimation and confidence intervals from complete, singly censored and multiply censored data. Simultaneous confidence bands. Uncertain censoring times. Quantities of interest in reliability applications. Maximum likelihood for location-scale and log-location-scale distributions. Confidence intervals for functions of parameters. Comparison of confidence interval procedures. Reliability prediction. Concept of Software reliability, software reliability modeling and estimation.

Probability plotting and choosing a failure-time distribution: Linearizing location-scale-based distributions. Graphical goodness of fit. Probability plots with specified shape parameters. Application of probability plotting. Weibull probability plot. model selection and validation.

Failure-time regression analysis: Failure-time regression models. Scale-accelerated failure-time model. Quadratic regression model. Checking model assumptions. Product comparison: an indicator-variable regression model. The proportional Hazards failure-time model and its applications in reliability.

Accelerated test models and analyzing accelerated life test data: Accelerating variables. acceleration models. Guideline for the use of accelerating models. Non-parametric and graphical methods for presenting and analyzing accelerated life test (ALT) data. Likelihood methods for analyzing right-censored data from an ALT. Suggestions for drawing conclusions from ALT data. Potential pitfalls of accelerated life testing.

References:

1. Ansell, J. I. and Phillips, M. J. (1994). “Practical Methods for Reliability Data Analysis”, Oxford: Clarendon Press.
2. Bain, L. J. and Engelhardt, M. (1991). “Statistical Analysis of Reliability and Life Testing Models, Theory and Methods (Second Edition)”, New York: Marcel Dekker.
3. Balakrishnan, N. (Editor) (1995). “Recent Advances in Life-Testing and Reliability”, Boca Raton, FL: CRC Press.
4. Balakrishnan, N. and Rao, C. R. (Eds.) (2001). “Handbook of Statistics, Volume 20: Advances in Reliability”, Elsevier Science B. V., The Netherlands.
5. Barlow, R. E. and Proschan, F. (1975). “Statistical Theory of Reliability and Life Testing”, New York: Holt, Rinehart and Winston.
6. Blischke, W. R. and Murthy, D. N. P. (Editors) (1996). “Product Warranty Handbook”, New York: Marcel Dekker.
7. Crowder, M. J., Kimber, A. C., Smith, R. L., and Sweeting, T. J. (1991). “Statistical Analysis of Reliability Data”, New York: Chapman & Hall.
8. Kalbfleisch, J. D. and Prentice, R. L. (1980). “The Statistical Analysis of Failure Time Data”, New York: John Wiley & Sons.
9. Kozlov, B. A. and Ushakov, I. A. (1970). “Reliability Handbook”, New York: Holt, Rinehart, and Winston.
10. Krishnaiah, P. R. and Rao, C. R. (Eds.) (1998). “Handbook of Statistics, Volume 7: Quality Control and Reliability”, Elsevier Science B. V.
11. Kume, H. (1985). “Statistical Methods for Quality Improvement”, The Association for Overseas Technical Scholarship: Japan.
12. Lawless, J. F. (1982). “Statistical Models and Methods for Lifetime Data”, New York: John Wiley & Sons.
13. Meeker, W. Q. and Escobar, L. A. (1998). “Statistical Methods for Reliability Data”, New York: John Wiley & Sons.
14. Nelson, W. (1990). “Accelerated Testing: Statistical Models, Test Plans, and Data Analyses”, New York: John Wiley & Sons.
15. Tobias, P. A. and Trindade, D. C. (1995). “Applied Reliability (Second Edition)”, New York: Van Nostrand Reinhold.

STAT M13: Actuarial Statistics and Data Mining

(Full Marks: 100 (Part A 40 Marks & Part B 60 Marks),

2 Lectures per week, Credit: 3.5)

(Year Final Exam: 75 Marks, 5 Questions out of 8 are to be answered)

Part A

Actuarial Statistics:

Definition of actuarial science, It's relationship with life insurance, Important uses of actuarial statistics specially in the context of Bangladesh.

The theory of rates of interest and discount including the theoretical continuous case of forces of interest and discount, Annuities and sinking funds, Including the continuous case, Practical and theoretical applications primarily to mortgages and bonds, Yield rates.

The economics of insurance, Utility theory, Application of probability to problems of life and death, The determination of single premiums for insurances and annuities in both the discrete and continuous case.

. Theory and practice of pension plan funding, Assumptions, Basic actuarial functions and population theory applied to private pensions.

Survival distributions and life tables, Life insurance, Life annuities, Net premium, Premium series, Multiple life functions, Multiple decrement models, Valuation theory for pension plans, The expense function and dividends.

References:

1. Kellison, S. G. (1970): Theory of Interest, Richard D Irwin
2. Bowers, N. L. et. Al. (1978): *Actuarial Mathematics*, Elasea
3. Frank Ayres, Jr., *Mathematics of Finance*
4. Bathen, R. W., *Mortality Table Construction*
5. Jordan, C. W. Jr. (1952): *Life Contingencies,* Society of Actuaries
6. Micheal M. Parmenter (1988): *Theory of Interest Life Contingencies with Premium* ACTEX

Part B

Data Mining:

Overview: Meaning of Data Mining and Knowledge Discovery basic, Data Mining tasks: Classification, Regression, Time Series Analysis, Prediction, Clustering, Summarization, Association, Rules, Sequence Discovery. Development of Data Mining, Data Mining issue and Mining Metrics, Social Implications of Data Mining.

Related Concepts of Fuzzy Sets: Introduction, classical sets, Set operation, Boolean logic, Basic concepts of Fuzzy sets, Other representations of Fuzzy sets, Determination of Member ship Functions, Fuzzy sets Properties, Operations on Fuzzy sets, Logic Operations, Algebraic Operations on Fuzzy sets.

Fuzzy Relations: Classical Relations, Classical Reasoning, Fundamentals of Fuzzy Relations, Operations on Binary Fuzzy Relations, Types of Fuzzy Relations, Fuzzy Reasoning Concluding Remarks, Bibliography, Web Resources. Data base/OLTP Systems, Logic, Information Retrieval, Decision Support Systems, Dimensional Modeling, Multidimensional Schemas, Indexing, Data Warehousing, OLAP, Web Search Engines, Statistics, Machine Learning, Pattern Matching.

Data Mining Techniques: Statistical Perspective on Data Mining: Point Estimation, Models Based on Summarization, Bayes Theorem, Hypothesis Testing, Regression and Correlation. Similarity Measures; Decision Tree; Genetic Algorithms.

Neural Networks: Background, NN Supervised Learning, The basic Neuron Model, The perception, The Multiplayer Perception Recurrent Network, Hopfield Networks, Boltzanann Machine, Kohonen Self-Organizing Map: Background, Description, Determining the winning neuron, Learning algorithm.

Classification: Issues in Classification, Statistical-Based Algorithms, Regression, Bayesian Classification, Distance-Based Algorithms, K. Nearest Neighbors, Decision Treee-Based Algorithms, ID3, C4.5 and C5.0, CART, Neural Network-Based Algorithms, Propagation, Radial Basis Function Networks, Perceptrons, Rule-Based Algorithms, Generating Rules from a DT, Generating Rules from a Neural Net.

Clustering: Similarity and Distance Measures, Outliers, Hierarchical Algorithms, Agglomerative Algorithms, Divisive Clustering, Partitional Algorithms, Minimum Spanning Tree, Squared Error Clustering Algorithm, K-Means Clustering, Nearest Neighbor Algorithm, PAM Algorithm, Bond Energy Algorithm, Clustering with Genetic Algorithms, Clustering with Neural Networks, Clustering Large Databases, Clustering with Categorical Attributes.

Association Rules: Meaning of Association Large Itemsets, Basic Algorithms, Apriori Algorithm, Sampling Algorithm, Partitioning, Parallel and Distributed Algorithms, Data Parallelism, Task Parallelism, Advanced Association Rule Techniques, Quantitative Association Rules, Correlation Rules, Measuring the Quality of Rules.

Web Mining: Web Content Mining, Crawlers, Harvest System, Virtual Web View, Personalization, Web Structure Mining, PageRank, Clever, Web Usage Mining, Preprocessing, Data SDtructures, Pattern Discovery, Pattern Analysis.

References:

1. Dunham, M. H. (2003): *Data Mining*, Pearson Education Asia.
2. Ibrahim, A. M. (2004): *Fuzzy Logic for Embedded Systems Applications*, Elsevier Science, USA.
3. Larose, D.T. (2006): *Data Mining-Methods and Models*, Wiley-Interscience, India.
4. Schalkoff, R. (2005): *Pattern Recognition- Statistical*, Structural and Neural Approaches, John Wiley & Sons, N. Y.

STAT M14: Experimental Design

(Full Marks: 100, 2 Lectures per week, Credit: 3.5)

(Year Final Exam: 75 Marks, 5 Questions out of 8 are to be answered)

Linear Estimation, Linear Hypothesis, General Principles and Practical Difficulties in Controlled Experiments, Strategies in Planning Experimental Programs.

S^n Factorial Experiments and Their Analysis, Confounding, Fractional Replications with Their Construction and Analysis, Split-Split-Plot Design, Split Block Design, Repeated Measurement Design, Incomplete Block Design, Balanced Incomplete Block Design (Inter and Intra-Block Analysis).

Missing Plot, Galois Field and Finite Projective Geometry and Its Application, Construction of Some Commonly Used Balanced Incomplete Block (BIB) Design, Orthogonal Latin Squares Design, Lattice Design, Youden Square Design, Partially Balanced Incomplete Block Design, Change-Over Design, Groups of Experiments, Optimality of Design.

Text:

1.Montegomery, D. C. (2003): *Design and Analysis of Experiments*, 5th edition, John Wiley, New York.

References:

1. Das, M. N. and Giri, N. C. (1997): *Design and Analysis of Experiments*, 2nd edition, New International (P) Ltd., India
2 .Kempthorne, O. (1952): *The Design and Analysis of Experiments*, Wiley, New York.
3 .Graybill, F. A. (1961): *An Introduction to Linear Statistical Models*, Vol. I, McGrow-Hill, New York.
4. Federer, W. T. (1955): *Experimental Design: Theory and Application*, Macmillan, New York.
5. Cochran, W. G. and Cox, D. R. (2000): *Experimental Design,* 2nd edition, John Wiley, New Delhi.
6. John, P. W. M. (1971): *Statistical Design and Analysis of Experiments,* Wiley, New York.

Practical/Field Works/Project and Viva-Voce fore the General Group –A : (Non this)

STAT M15: Lab I

Course Title: Statistical Data Analysis-I

(Full Marks: 100, Credit: 3.0)

Advanced Statistical Inference (Part A, Marks: 50):

Non-parametric Estimation: Estimation of the density curve and regression curve of real life data using smoothing technique (Kernel method and Fourier series method), finding the optimal bandwidth using argmin method.

EM Algorithm: Estimation of maximum likelihood estimators in the present of missing observation using Expectation-Minimization algorithm.

Simultaneous Confidence Interval: Construction of simultaneous confidence interval specifically in the case of multiple comparisons in one way ANOVA model using Bonferroni's method and scheffe's method.

Sobel Test: Testing the significance of mediation effect on the dependent variable using Sobel test according to Baron and Kenny approach.

Advanced Multivariate Analysis and Categorical Data Analysis (Part B, Marks: 50):

1. Checking normality of a multivariate data set and transforming a multivariate data set into normal one.

2. Constructing 95% T^2 simultaneous and Bonferroni confidence intervals, and determining lengths and directions for the axes of the 95% confidence ellipsoid for mean.

3. Determining sample principal components and fining proportions of the total sample variance explained by different components individually.

4. Finding important factors using different approaches, applying different rotation methods and finding factor scores using different methods.

5. Determining sample canonical correlations, constructing canonical variates, and preparing table showing the canonical variate coefficients and the sample correlations of the canonical variates with their component variables.

6. Clustering by using different linkage methods, constructing dendrograms, and a Kmeans clustering program.

STAT M16: Lab II

Course Title: Statistical Data Analysis-II

(Full Marks: 100, Credit: 3.0)

Econometric Analysis & Environmental Statistics (Part A, Marks: 50):

1. Fitting a multiple regression model in economic data, applying different post estimation tests and designing Chow test for checking structural change of a relationship during a specified period.

2. Fitting dummy variable regression models, deseasonalizing given time series data using dummy variable method and checking structural change of a regression during a period using dummy variable method.

3. Identifying multicollinearity problem and its consequences in a given economic time series data, and solving the multicollinearity problem using different methods.

4. Checking the presence of autocorrelation and heteroskadasticity in data applying different appropriate methods and taking appropriate remedial measures for the problems in fitting time series regression.

5. Carry out RESET and LM tests of specification errors, and identifying the type and consequences of specification error.

6. Fitting recursive least squares to check the stability of estimated regression coefficients during a specified period.

7. Fitting nonlinear regression model, e.g. growth models and Cobb-Douglas production function.

8. Preparing a panel data set, and fitting pooled regression, fixed effect and random effects models to the data.

Actuarial Statistics (Part B, Marks: 25):

Computing Compound Interest and Discount: Problems of compound interest on banking and insurance, determining forces of interest, estimating discount-nominal rate of discount & nominal rate of interest.

Construction of Amortization Schedule: Constructing of amortization schedule & bond amortization schedule; bond, yield rate & book value computing.

Actuarial Science: Application of probability to problems of life and death, pension planning, design of insurance; life table for estimating pension, life insurance & life annuities and multiple life functions related problems.

Data Mining (Part C, Marks: 25):

Determination of partition of set of data by sum of squares of errors, Clustering criteria, Hierarchical clustering by Nearest Neighbor, Further Neighbor, K-means method or algorithm, Determination of minimum distance decision boundary, Performing K-NN classification using the Euclidean and statistical Matrix. Application of different rules of data mining, Classification by regression tree, decision tree, Bayesian approach and neural network.

STAT M17: Lab III

Course Title: Statistical Data Analysis-III

(Full Marks: 100, Credit: 3.0)

Sample Survey (Part A, Marks: 50):

1. Estimation of population mean, total and estimation of standard errors of the estimator of the following sampling procedures: (i) Multistage sampling (ii) Multiphase sampling (iii) Adaptive cluster sampling (iv) Stratified adaptive cluster sampling (v) Systematic and strip adaptive cluster sampling.

2. Determination of sample size under different sampling techniques by using various techniques and estimating population size using direct inverse sampling.

3. Considering response class and non response class finding expected cost of a survey, estimating expected number of enumerator in a survey.

4. Fining mean square prediction error of spatial sampling.

Population Studies (Part B, Marks: 50):

1. Construction of single and multiple decrement tables.

2. Estimation of parameters in couple year protection.

3. Estimation of indices of Bongaart’s model and fitting this model.

4. Smoothing of age-specific fertility rates generated by Gompertz relational fertility model.

STAT M18: Project
(Full Marks: 100, Credit: 3.0)

STAT M20: Viva-Voce
(Full Marks: 100, Credit: 2.5)

Comprehensive Practical Viva-Voce ,Thiess and presentation for the theses Group(Group - B)

STAT M19: Comprehensive Practical
(Full Marks: 100, Credit: 4.0)

STAT M20: Viva-Voce
(Full Marks: 100, Credit: 2.5)

STAT M21: Dissertation
(Full Marks: 200, Credit: 5.0)

STAT M22: Presentation of the Dissertation
(Full Marks: 100, Credit: 3.0)

Panel of Paper setters, Examiners, and Tabulators

1. All teachers, Dept. of Statistics, Islamic University, Kushtia.
2. All teachers, Dept. of Mathematics, Islamic University, Kushtia
3. All teachers, Dept. of Computer Science & Engineering, Islamic University, Kushtia.
4. All teachers, Dept. of Economics, Islamic University, Kushtia
5. All teachers, Dept. of Statistics, R. U., Rajshahi.
6. All teachers, Dept. of Applied Mathematics, R. U., Rajshahi.
7. All teachers, Dept. of Mathematics, D. U., Dhaka.
8. All teachers, Dept. of Statistics, Biostatistics and Informatics, D. U., Dhaka
9. All teachers, Dept. of Applied Statistics, D. U., Dhaka.
10. All teachers, Dept. of Mathematics, C. U., Chittagong.
11. All teachers, Dept. of Statistics, C. U., Chittagong.
12. All teachers, Dept. of Mathematics, J. U., Savar, Dhaka.
13. All teachers, Dept. of Statistics, J. U., Savar, Dhaka.
14. All teachers, Dept. of Statistics Discipline, K. U., Khulna.
15. All teachers, Dept. of Mathematics, SUST, Sylhet.
16. All teachers, Dept. of Statistics, SUST, Sylhet.
17. All teachers, Dept. Statistics, BRU, Rangpur.
18. All teachers, Dept. of Statistics, BAU, Mymensing.
19. All teachers, Dept. of Statistics, CoU, Comilla.
20. All teachers, Dept. of Statistics, JnU, Dhaka.
21. All teachers, Dept. of Mathematics, JnU, Dhaka.
22. All teachers, Dept. of Mathematics, PSTU, Pabna.
23. All teachers, Dept. of Statistics, BSMRSTU, Gopalgonj.
24. All teachers, Dept. of Statistics, MBSTU, Tangail.
25. All teachers, Dept. of Statistics, HDSTU, Dinajpur.

(Altaf Hossain)
Chairman
Department of Statistics
Islamic University, Kushtia, Bangladesh.